In vivo RNA structural probing of uracil and guanine base pairing by 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC)

David Mitchell III[1,2], Andrew J. Renda[2,3], Catherine A. Douds[1,2,3], Paul Babitzke[2,3], Sarah M. Assmann[4], and Philip C. Bevilacqua[1,2,3*]

[1]*Department of Chemistry, The Pennsylvania State University, University Park, PA 16802, USA.*
[2]*Center for RNA Molecular Biology, The Pennsylvania State University, University Park, PA 16802, USA.*
[3]*Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802, USA.*
[4]*Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA.*

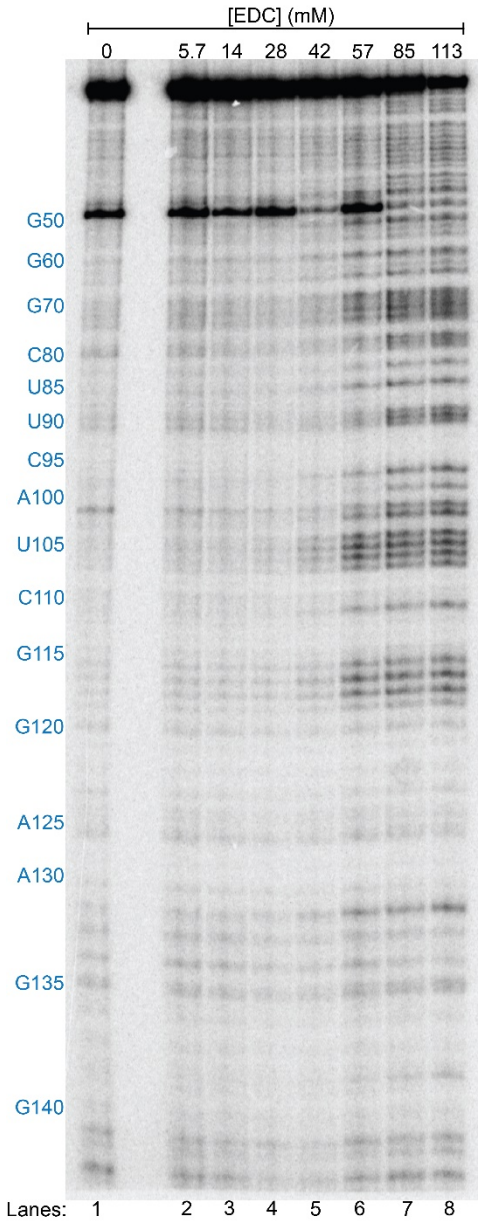
**Corresponding author*: e-mail: pcb5@psu.edu; Phone: (814) 863-3812; Fax: (814) 865-2927

Running Title: In vivo probing of uridine base pairing

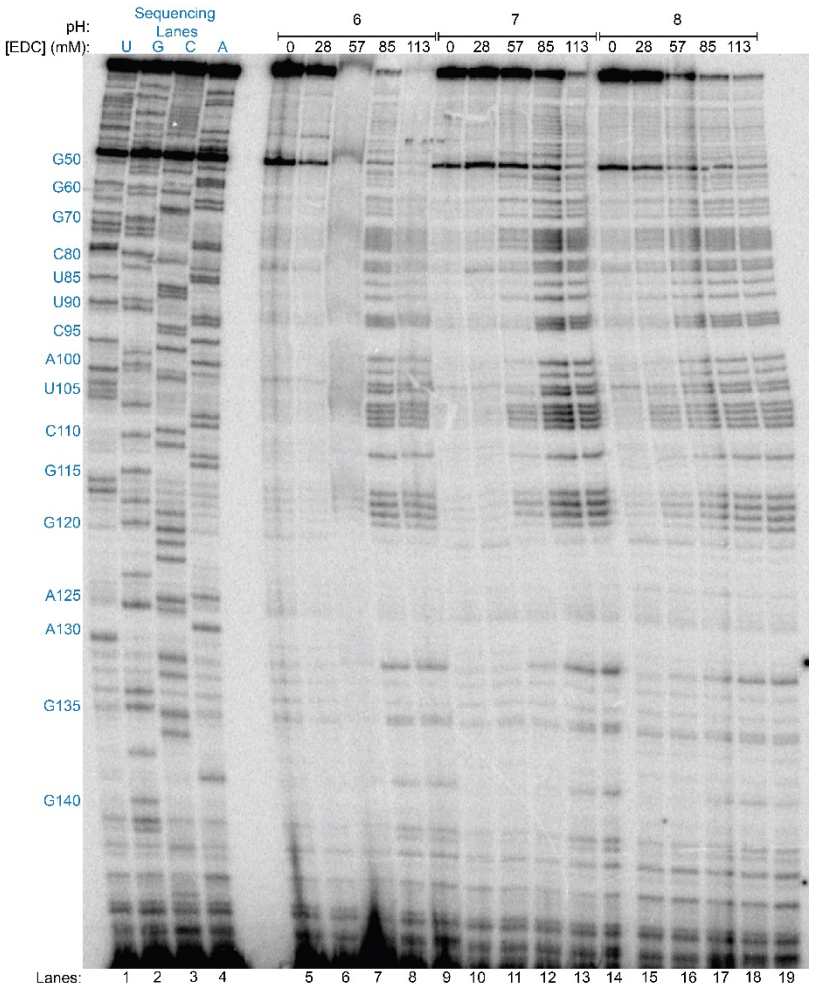
Keywords: EDC, N-(3-dimethylaminopropyl)-N'-(ethylcarbodiimide), carbodiimide, RNA structure, in vivo RNA probing

Supplementary Text

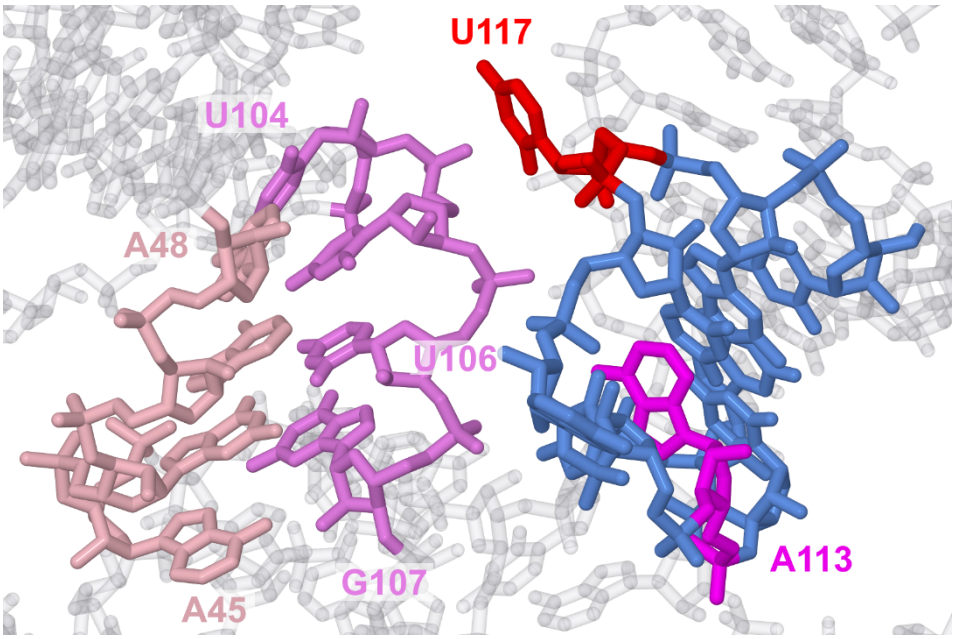
EDC Reaction Quench: We quenched the EDC reaction with rice seedlings by a three-step process. First, we added 1 g of solid dithiothreitol (DTT) prior to three water washes of the plant tissue. Tests showed that DTT prevents EDC from reacting with uracils or guanines *in vitro* (Supplemental Figure S4A). Second, after the water washes, the tissue was quickly frozen in liquid N_2. Third, the sample was thawed in a lysis buffer containing additional DTT at 50 mM. The reaction was adequately quenched by this three-step process as revealed by time points for reactivity of various nucleotides that extrapolated back to the origin (Supplemental Figure S4B), as well as by a quench control. In the quench control, we doped the 57 nt ATP aptamer RNA, which is not natural to rice and thus contains a sequence not found in total rice RNA, into the lysis buffer used in the RNA extraction. We found that there was no EDC-specific reaction of the ATP aptamer (Supplemental Figure S4C), indicating that the EDC had been successfully quenched by the prior treatment. Importantly, RT extension only occurs when the ATP aptamer is present (Supplemental Figure S4C, Lanes 5, 7, 10, 12).



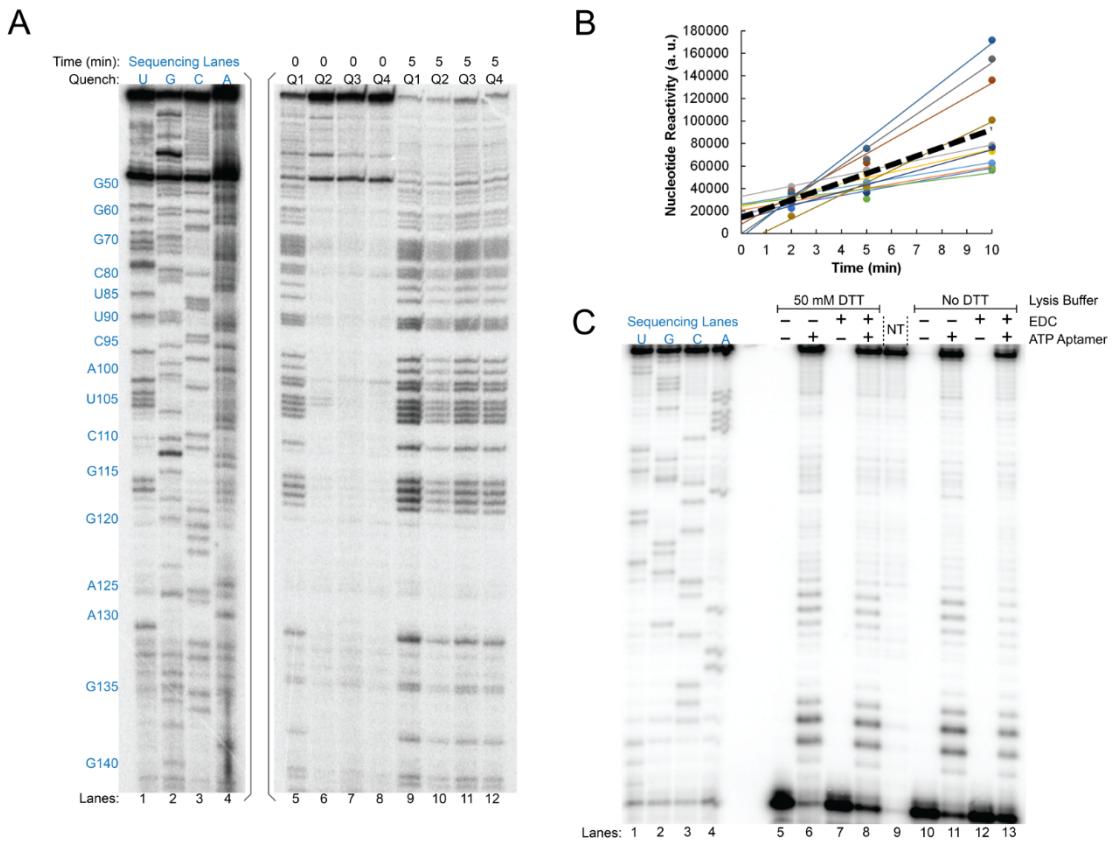
Supplemental Figure S1. *In vitro* modification of rice 5.8S rRNA by EDC for 2 min reaction duration analyzed by denaturing page of cDNAs after reverse transcription. Reactions with the indicated EDC concentrations. A control reaction lacking EDC and reactions with 5.7 mM to 113 mM EDC are shown. Blue text to the left indicates the sequence of the examined range of G53 to C143.



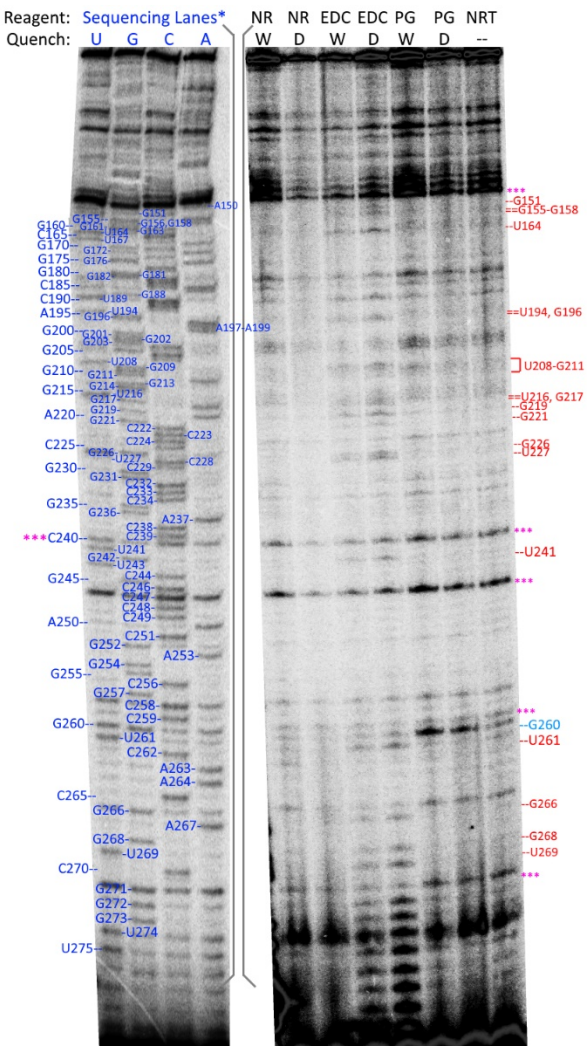
Supplemental Figure S2. *In vitro* EDC modification of rice 5.8S rRNA *in vitro* at various pH and EDC concentrations. Denaturing PAGE analysis of cDNAs generated after reverse transcription. Reaction conditions at pH 6, pH 7, and pH 8 are shown along with dideoxy sequencing lanes.



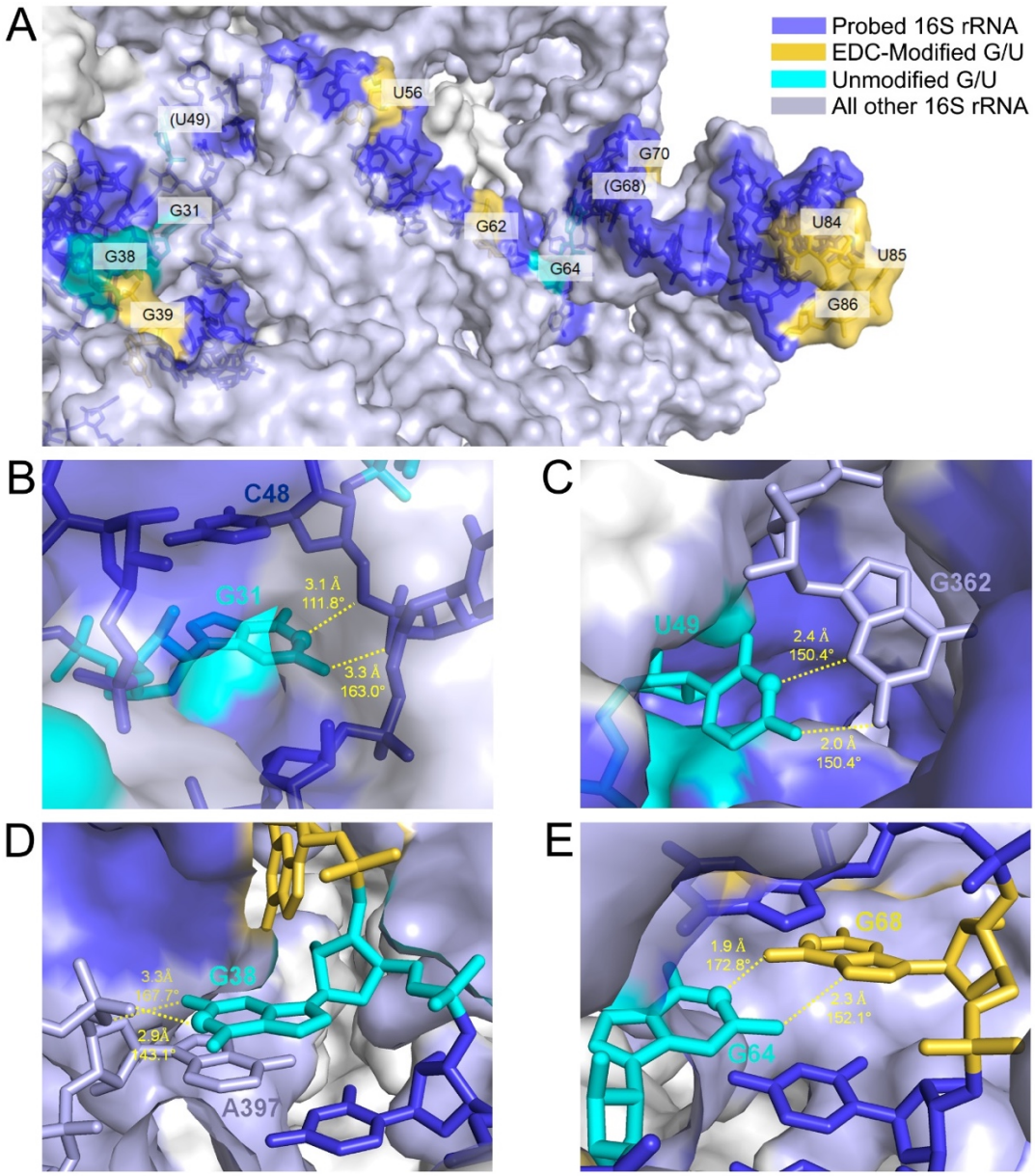
Supplemental Figure S3. Cryo-EM structure of *Saccharomyces cerevisiae* 60S subunit (PDB: 5GAK), a homolog of rice 60S subunit, is used here as no rice ribosome structure currently exists. Shown exclusively is 5.8S rRNA. The long-range helix at left shows A45 to A48 in pink and U104 to G107 in violet. Note that G107 is in a sheared base pair and U106 forms a wobble pair. The stem-loop from G111 to G119 is blue, with the splayed out U117 in red and A113 in magenta. This stem-loop has an identical sequence in rice. The remainder of 5.8S rRNA is shown in transparent white.



Supplemental Figure S4. *In vitro* probing of rice 5.8S rRNA by EDC to test quench conditions. (*A*) Tests of DTT and sodium acetate reaction quenches analyzed by denaturing PAGE of cDNAs after reverse transcription. The dideoxy sequencing lanes at left were run on a different part of the same gel, and the transposition of these lanes is indicated by the grey brackets. Four different quench compositions were examined: water (Q1), 2.5 mM DTT (Q2), 1 M sodium acetate, pH 5 (Q3), and a combination of 1.3 M DTT and 1 M sodium acetate, pH 5 (Q4). Times are when the quench solution was added, with 0 min indicating addition of the quench before adding 113 mM EDC and 5 min indicating addition of the quench 5 min after reacting total rice RNA with EDC. (*B*) Plot of normalized nucleotide reactivities against reaction time for EDC-modified nucleotides between U102 and U131. Lines represent linear fits. The bold line indicates the fit to the average reactivity for all examined nucleotides. (*C*) Test of lysis buffer composition analyzed by denaturing page of cDNAs after reverse transcription. The sequencing lanes are for ATP aptamer RNA, an RNA sequence not found in rice total RNA, which was doped into lysis buffer before RNA extraction for lanes 6, 8, 11, and 13. Lanes 5, 7, 10, and 12 do not contain ATP aptamer RNA. Lane 9, labeled NT, contains untreated ATP aptamer RNA not added to lysis buffer for which reverse transcription was done separately. Less RNA was added to the RT reaction for NT, which accounts for the lower band intensity in lane 9 compared to lanes 6, 8, 11, and 13.



Supplemental Figure S5. Comparison of *in vivo* EDC and phenylglyoxal modification of rice 28S rRNA analyzed by denaturing PAGE of cDNAs after reverse transcription. Specified here is the range from A150 to C270. EDC and phenylglyoxal (PG) modifications under conditions where either a water wash (W) or 1 g of DTT (D) was used as a reaction quench are shown, along with dideoxy sequencing lanes. The dideoxy sequencing reactions were performed separately and run on a separate gel, as indicated by the grey brackets and asterisk in the text next to Sequencing Lanes. Rice tissue not treated with reagent nor subjected to quenching is shown as NRT. Dark blue text at left indicates the sequence of 28S rRNA. Red text at right indicate nucleotides modified by EDC. Light blue text at right indicate nucleotides modified by both EDC and phenylglyoxal. Magenta asterisks indicate natural reverse transcription stops.



Supplemental Figure S6. Crystal structure of the *Escherichia coli* 70S ribosome (PDB: 4V9D) to show uracils (U) and guanines (G) within the examined range for EDC reactivity. Lack of reactivity of some Gs and Us can be explained by solvent inaccessibility and hydrogen bonding, while others can be explained by hydrogen bonding alone. (*A*) Comparison of EDC-modified and EDC-unmodified Gs and Us within 16S rRNA. In this and all subsequent panels, the examined range (1-90) within 16S rRNA is in dark blue, the remainder of 16S rRNA is in pale blue, Us and Gs modified by EDC (see Figure 6 in the main text) are in gold, Us and Gs unmodified by EDC are in cyan, and other ribosomal RNAs are in grey. (*B*) G31 is partially buried and in position to form a hydrogen bond between its N1 and the bridging O5' of C48. The N1G or N3U is shown as a sphere in this and subsequent panels. (*C*) U49 is also buried within the ribosome. A slice of the ribosome structure is removed to allow easy viewing. U49 forms a sugar edge interaction with G362. (*D*) G38 is in position to form a hydrogen bond between N1 and a non-bridging phosphate oxygen of A397. (*E*) G64 forms a Hoogsteen base pair with G68, which in turn forms a sheared pair with A101 (not shown).